An autonomous and reversible Maxwell's demon

A. C. Barato and U. Seifert

II. Institut für Theoretische Physik, Universität Stuttgart - Stuttgart 70550, Germany

Abstract – Building on a model introduced by Mandal and Jarzynski [Proc. Natl. Acad. Sci.
U. S. A., **109**, (2012) 11641], we present a simple version of an autonomous reversible Maxwell's
demon. By changing the entropy of a tape consisting of a sequence of bits passing through
the demon, the demon can lift a mass using the coupling to a heat bath. Our model becomes
reversible by allowing the tape to move in both directions. In this thermodynamically consistent
model, total entropy production consists of three terms one of which recovers the irreversible
limit studied by MJ. Our demon allows an interpretation in terms of an enzyme transporting and
transforming molecules between compartments. Moreover, both genuine equilibrium and a linear
response regime with corresponding Onsager coefficients are well defined. Efficiency and efficiency
at maximum power are calculated. In linear response, the latter is shown to be bounded by 1/2
if the demon operates as a machine and by 1/3 if it is operated as an eraser.

Maxwell's demon is a device which allows to extract work from a heat bath by rectifying fluctuations. It has inspired discussions on the foundations of thermodynamics for more than a century [1, 2]. Recently, its inherent aspect of information processing in a thermodynamic setting has turned into focus in theoretical studies of specific devices [3–14], in the derivation of fluctuation theorems and second law like inequalities modified for the presence of information [15–22], and through pioneering experimental implementations [23, 24]. Typically, these schemes require the action of a macroscopic observer and a *feedback* controller.

An explicit model for an *autonomous* demon has recently been introduced by Mandal and Jarzynski (MJ) [14] for which the relation between mechanical work and information processing, including the notoriously subtle issue of the cost of measurement and/or erasure, becomes particular transparent. In the MJ model, a frictionless tape composed of bits passes through the demon, which is also in contact with a thermal bath and a mechanical reservoir represented by a mass that can go up or down. During a certain time interval the demon interacts with the bit, jumping between its internal states with transition rates that depend on the mass. Between two interaction time-intervals and ignoring correlations between the bits on the tape, a new incoming bit is instantaneously cho-

sen from the probability distribution corresponding to the ratio of 0's and 1's on the tape. The model reaches a periodic steady state, where the outgoing tape is a record of how the internal states of the demon are influenced by the mass. Remarkably, MJ [14] could show that this model fulfills a second law like inequality, with the entropy difference between the incoming and outgoing tape bounding the work that the demon delivers to lift the mass. Saturating this equality, however, is arguably not equivalent to equilibrium since the MJ model contains a fundamental asymmetry because the (frictionless) tape is fed into the demon from one direction, with no option of a spontaneous reversal of directions, as it would happen if the tape was in contact with a thermal reservoir as well.

The main purpose of this letter is to introduce a model for an autonomous reversible demon whose parameter space contains genuine thermodynamic equilibrium. Inspired by the MJ model, we reformulate and generalize it as a fully continuous-time stochastic process. Each transition, in principle, is reversible by allowing the tape to move in either direction and not just from left to right as in the MJ model. This system reaches a nonequilibrium steady state for which the thermodynamic entropy production is easily identified using the framework of stochastic thermodynamics [25]. Total entropy production has then three contributions, the first corresponding to the difference be-
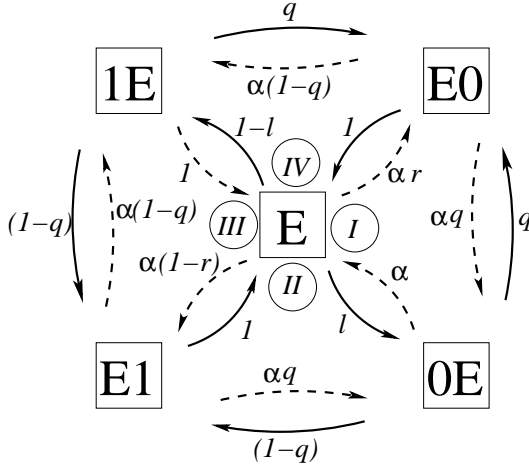
Fig. 1: The five states model with its transition rates. The full lines represent the transitions where the tape moves from left to right, whereas the dotted line represent the reverse transitions. All rates in the figure should be multiplied by k, which sets the times scale of the transitions. The cycles I, II, III, and IV, as defined in equation (1), are indicated as well.
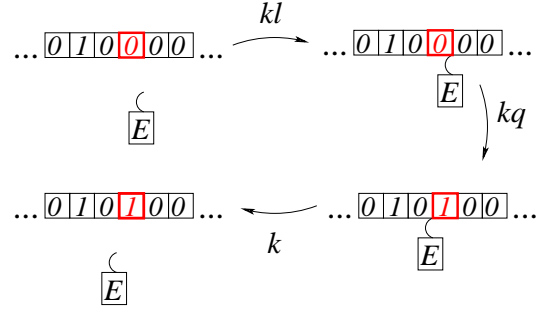

Fig. 2: The moving demon picture. First the demon binds to a bit 0 with rate kl, it then transforms the 0 into a 1 moving to the right with rate $k(1-q)$, and, finally, it detaches from the tape with a rate k. Therefore, we have a representation of cycle C_{II} defined in (1).

tween the entropy change of the tape and mechanical work as identified by MJ, the second one arising from the fact that the tape can reenter from the right side, and the third one due to the friction associated with the motion of the tape.

Genuine equilibrium being part of the phase space of our system, we can develop a linear response treatment with the corresponding identification of the Onsager coefficients. In our analysis, we will exploit an alternative chemical representation of the demon as an enzyme that transports and transforms particles from one reservoir to another. Another mapping of a demon onto a chemical model has been suggested independently recently in a somewhat related model [26].

The network of five states and the transition rates between them define our model as shown in Fig. 1. First, we consider the case where the tape moves only from left to right, which means that only the transitions represented by the full lines in Fig. 1 are possible. In the central state E the demon is not interacting with a bit from the tape. From E the system can either jump to $0E$ or $1E$, with rates kl and $k(1-l)$, respectively, where k sets a time scale. These jumps represent a bit coming from the left to interact with the demon. Therefore, the parameter l is interpreted as the probability of an incoming 0, while $1-l$ gives the probability of an incoming 1.

If the system is at $0E$ or $1E$ it either jumps to $E0$ or $E1$ corresponding to the state of the bit going out to the right. The transition between the incoming and the outgoing bit arises from an interaction between the demon and the bit, which may transform the incoming bit and lift (or lower) the mass, through an auxiliary process involving two auxiliary states E_0 and E_1. An auxiliary transition from E_0 to E_1 (from E_1 to E_0) involves lifting (lowering)

the mass by an overall distance h. The transitions between these two auxiliary states happen on a time scale much faster than $1/k$. Hence, they are equilibrated on the latter scale leading to a stationary probability q for being in E_0 and $1-q$ for being in E_1, where $q = e^f/(1+e^f)$, with $f \equiv mgh/k_BT$ and T is the temperature of the heat bath. While the initial state E_b ($b = 0, 1$) for this auxiliary process is given by the state bE characterizing the incoming bit, the final state for this auxiliary process $E_{b'}$ ($b' = 0, 1$) determines the state of the outgoing bit Eb'. The sole purpose of this auxiliary two-state process is to yield effective rates for the main network shown in Fig. 1. We note that this scheme corresponds to the limit where the interaction time becomes infinite in the MJ model. Moreover, if the transition goes from $0E$ ($1E$) to $E1$ ($E0$) then the mass is lifted (lowered) by a height h and a quantity of work k_BTf (we set in the following $k_BT = 1$) will be delivered to (taken from) the mechanical reservoir. On the other hand, if the transition went from $0E$ to $E0$ or from $1E$ to $E1$, then there is no mechanical work involved.

From the states $E0$ and $E1$ the system jumps to the state E with rate k. In these transitions, the demon let the transformed bit leave to the right in the outgoing tape and comes back to the central state. Therefore, in a complete cycle starting and finishing at the central state E, the demon gets a bit from the incoming tape, interacts with this bit, possibly changing it, and then sends the final bit to the outgoing tape.

We can also consider the model as a demon walking along the tape and changing its bits, see Fig. 2. In this case the demon binds to the tape, interact with the bit, then unbinds from the tape, and proceeds by repeating the same process with the next bit.

Summarizing, the model has four cycles, which are

$$\mathcal{C}_I \equiv (E, 0E, E0, E), \qquad \mathcal{C}_{II} \equiv (E, 0E, E1, E),$$
$$\mathcal{C}_{III} \equiv (E, 1E, E1, E), \qquad \mathcal{C}_{IV} \equiv (E, 1E, E0, E). \quad (1)$$

In the cycle $\mathcal{C}_{II}$ an incoming 0 is transformed into an outgoing 1, therefore the demon delivers a quantity of work

f to the mechanical reservoir. On the other hand, in the cycle $\mathcal{C}_{IV}$ an incoming 1 is transformed into an outgoing 0 and the demon receives an energy f from the mechanical reservoir. In the other two cycles the incoming bit is not changed and no work is exchanged with the mechanical reservoir. The average time for the completion of a cycle is the same for all cycles and given by $3k^{-1}$, which follows from the fact that the escape rate of all states is equal to k. Moreover, the four cycles are the only four possible paths that start and finish at E, and the probability of going through cycle $\mathcal{C}_{II}$ is $l(1-q)$ and through cycle $\mathcal{C}_{IV}$ is $(1-l)q$. Therefore, the rate of work flowing to the mechanical reservoir in the stationary state is

$$\dot{w}_m = kf(l-q)/3. \tag{2}$$

The Shannon entropy of the incoming tape is given by

$$H(l) \equiv -l\ln l - (1-l)\ln(1-l). \tag{3}$$

This quantity measures how much information the incoming tape contains. Moreover, it is easy to calculate the stationary state probability distribution of the model (see equation (8) below), which gives $P_{E0}/P_{E1} = q/(1-q)$, where P_i denotes the stationary probability distribution of state i. Hence, in the stationary state the entropy of the outgoing tape is $H(q)$ and the rate of information $\dot{h}$ is given by the entropy difference between the outgoing tape and the incoming tape divided by the average time to complete a cycle, leading to

$$\dot{h} = k\Delta H/3, \tag{4}$$

where $\Delta H \equiv H(q) - H(l)$.

Remarkably, the rates of mechanical work and information flow follow the inequality

$$\dot{h} - \dot{w}_m = k(\Delta H - W_m)/3 = kD(l||q)/3 \geq 0, \tag{5}$$

where $W_m \equiv (l-q)\ln(q/(1-q))$ is the average delivered work per cycle and $D(l||q) \equiv l\ln(l/q) + (1-l)\ln((1-l)/(1-q))$ is the Kullback-Leibler distance between the distribution of incoming and outgoing bits, which is known to be non-negative [27]. The expressions obtained for ΔH and W_m are identical to the expressions that are obtained in the MJ model [14] in the limit where the interaction time interval is large. As a first result, we have thus shown that the difference $\Delta H - W_m$ is given by $D(l||q)$. We note that in our model the transitions involving a bit coming from the left and the bit leaving to the right are modeled as a Poisson process with a finite waiting time, whereas in the MJ model these processes are instantaneous.

Depending on the value of l and q, the inequality (5) can bound the amount of mechanical work that can be extracted by transferring entropy to the tape or the amount of entropy reduction in the tape generated by the mechanical energy received from the falling mass [14]. Specifically, without loss of generality we consider $l \geq 1/2$, corresponding to an excess of 0 on the incoming tape. For $q > 1/2$

and $l > q$ we have $\dot{h} > 0$ and $\dot{w}_m > 0$, hence the demon is a machine that lifts a mass by transferring entropy to the tape. On the other hand, for $q > 1/2$ and $l < q$ both quantities change sign and the demon is an eraser that uses the energy from the falling mass to erase information (reduce entropy) of the tape. For $q < 1/2$ and $l < 1-q$ the demon also operates as an eraser, while for $q < 1/2$ and $l > 1-q$ we have $\dot{h} > 0$ and $\dot{w}_m < 0$, meaning that the demon just dissipates energy and does not perform any useful operation.

The above model is a fair description of a frictionless tape interacting with a Maxwell's demon that can lift a mass by exchanging entropy with the tape. However, all the reversed transition rates, represented by the dotted lines in Fig. 1, are zero and, therefore, entropy production in the sense of stochastic thermodynamics cannot be defined [25]. From this perspective, it is then somewhat surprising that the fluxes $\dot{h}$ and $\dot{w}_m$ still fulfill the second law like inequality (5). In the following, we consider all transition rates in Fig. 1 to be non-zero, which means that the tape can also move from right to left. Such a model has a well defined equilibrium. We would like to compare the standard entropy production with the dissipation rate in equation (5). Before considering the full model as a demon interacting with a moving tape we show that the model can also be seen as an enzyme that mediates chemical reactions between different molecules.

The chemical model can be described as follows. Two reservoirs of particles are separated by a membrane. In the left (right) reservoir molecules 0_L (0_R) and 1_L (1_R) are dissolved in a solvent. These molecules can bind to an enzyme E sitting on the membrane that transforms a molecule of one reservoir into a molecule of the other one. The concentration c_i of the species on the left [right] reservoir are cl [cr] for the substance 0_L [0_R] and $c(1-l)$ [$c(1-r)$] for the substance 1_L [1_R]. Moreover, the concentrations are small ($c \ll 1$) so that the chemical potentials $\mu_{0_L}, \mu_{1_L}, \mu_{0_R}$ and μ_{1_R} are related to the concentrations by $\mu_i = \ln(2c_i/c)$, where the additive constant $\ln 2$ is chosen for later convenience.

In the cycle $\mathcal{C}_I$, for example, first a molecule 0_L from the left reservoir binds to the enzyme, then the enzyme transforms it into a 0_R and, finally, the 0_R molecule is released into the right reservoir. In the reversed cycle $\overline{\mathcal{C}}_I$, a 0_R is taken from the right reservoir, transformed into a 0_L by the enzyme that is then released into the left reservoir. The parameter α, in Fig. 1, determines the rate of those transitions that correspond to taking a molecule from the right reservoir and releasing another molecule into the left reservoir: if $\alpha < 1$ the left to right transitions are faster than the reversed transitions. Furthermore, the enzyme E is connected to a mechanical reservoir, so that when it transforms a 0_L (0_R) into a 1_R (1_L), it delivers a quantity f of mechanical work to this reservoir. With the transition rates of Fig. 1 the escape rates of the four external states are the same, i.e., the average time the enzyme interacts with the molecules is the same for all four molecules.

Table 1: The affinities related to the respective transitions.

m	affinity $\mathcal{F}_m$	transition
1	$\ln \alpha^{-1}$	all transitions from left to right
2	μ_{0_L}	$E + 0_L \rightleftharpoons 0_L E$
3	μ_{1_L}	$E + 1_L \rightleftharpoons 1_L E$
4	μ_{0_R}	$E + 0_R \rightleftharpoons E0_R$
5	μ_{1_R}	$E + 1_R \rightleftharpoons E1_R$
6	$-f$	$0_L E \rightleftharpoons E1_R$ and $E0_R \rightleftharpoons 1_L E$

Table 2: The distances d_a^m entering equation (6), where a represents the cycles and m the individual affinities.

cycle	$\ln \alpha^{-1}$	μ_{0_L}	μ_{1_L}	μ_{0_R}	μ_{1_R}	$-f$
C_I	3	1	0	-1	0	0
C_{II}	3	1	0	0	-1	1
C_{III}	3	0	1	0	-1	0
C_{IV}	3	0	1	-1	0	-1

The model has the following six affinities (or thermodynamic forces): $\ln \alpha^{-1}$, related to the asymmetry between movements from left to right and from right to left, the four chemical potentials $\{\mu_i\}$, and $-f$. We define the stationary flux between states i and j as $J_{i,j} \equiv P_i w_{ij} - P_j w_{ji}$, where w_{ij} is the transition rate from i to j. Each affinity is related to a flux which enters the entropy production as $\dot{s} = \sum_m \mathcal{F}_m J_m$, where $\mathcal{F}_m$ denotes the affinity and J_m the flux between the pair of states related to the affinity. In Table 1, we show the affinities and the transitions they are related to.

Within stochastic thermodynamics, the entropy production can be written as cycle affinities multiplied by fluxes summed over the fundamental cycles as introduced in (1) [25]. In this formulation the entropy production takes the form

$$\dot{s} = \sum_{a=I}^{IV} J_a \sum_{m=1}^{6} d_a^m \mathcal{F}_m$$

$$= k(1-\alpha) \ln \alpha^{-1} + \sum_{a=I}^{IV} J_a \sum_{m=2}^{6} d_a^m \mathcal{F}_m. \quad (6)$$

where d_a^m is a generalized distance that indicates in how many of the transitions pertaining to the cycle a the affinity m appears (see Table 2). For later convenience, in the second equality we have separated the flux $J_1 = 1 - \alpha$ related to the affinity $\ln \alpha^{-1}$.

Using Table 2, the entropy production becomes

$$\dot{s} = k(1-\alpha) \ln \alpha^{-1}$$
$$+ J_{0E,E1}(\mu_{0_L} - \mu_{1_R} - f) + J_{1E,E0}(\mu_{1_L} - \mu_{0_R} + f)$$
$$+ J_{0E,E0}(\mu_{0_L} - \mu_{0_R}) + J_{1E,E1}(\mu_{1_L} - \mu_{1_R}). \quad (7)$$

For each external link a we can separate the current through it into two contributions, so that $J_a = J_a^+ - J_a^-$. If the link a connects states i and j (from i to j), then we have $J_a^+ = P_i w_{ij}$ and $J_a^- = P_j w_{ji}$. Moreover, given a link a that connects states bE and Eb' (with b, b' 0 or 1), calculating the stationary state fundamental fluxes we obtain

$$J_a^+ = k[X_a^+ + \alpha Y_a^+]/[3(1+\alpha)],$$
$$J_a^- = k\alpha[X_a^- + \alpha Y_a^-]/[3(1+\alpha)], \quad (8)$$

where

$$X_a^+ \equiv [l + b(1 - 2l)][q + b'(1 - 2q)],$$
$$Y_a^+ = X_a^- \equiv [q + b(1 - 2q)][q + b'(1 - 2q)],$$
$$Y_a^- \equiv [q + b(1 - 2q)][r + b'(1 - 2r)]. \quad (9)$$

With these definitions we can rewrite the entropy production in the instructive form

$$\frac{\dot{s}}{k} = (1-\alpha)\ln \alpha^{-1} + \frac{1}{3(1+\alpha)} \sum_{a=I}^{IV}(X_a^+ - \alpha^2 Y_a^-) \sum_{m=2}^{6} d_a^m \mathcal{F}_m, \quad (10)$$

which is useful in the following discussion.

Let us now go back to the picture where the demon interacts with the tape. Without loss of generality we consider $\alpha \leq 1$, which means that on average the tape moves from left to right. As a self-consistency condition, we have to impose that the ratio of probabilities of the outgoing bits to the right is equal the ratio of the incoming bits from the right. This leads to the condition

$$q/(1-q) = r/(1-r), \quad (11)$$

i.e., $r = q$. This means that the tape refed from the right has the same statistics as the outgoing tape to the right. With this self-consistency condition and using the definition (9), we obtain

$$\frac{1}{3(1+\alpha)} \sum_a X_a^+ \sum_{m=2}^{6} d_{am}\mathcal{F}_m = \frac{1}{3(1+\alpha)}(\Delta H - W_m), \quad (12)$$

as one contribution to the entropy production. This contribution is related to the difference between the statistics of the bits of the tape coming in from the left and the tape going out to the right. It is precisely the term we got in equation (5), obtained there for the case where $\alpha = 0$. Moreover, from (9) we also obtain

$$-\frac{1}{3(1+\alpha)}\alpha^2 \sum_a Y_a^- \sum_{m=2}^{6} d_{am}\mathcal{F}_m = \frac{1}{3(1+\alpha)}\alpha^2 D(q||l). \quad (13)$$

This term is due to the fact that the tape coming in from the right and tape going out to the left have different probability distributions. In our model, we do not keep correlations between the bits, considering the simplest approximation where a tape containing a 0 with probability q is

refed from the right to the left, where the probability of a 0 is l. Note that for small α, where transitions from right to left become rare, the term (13) is negligible compared to (12). Combining these terms, the entropy production of the tape model is given by

$$\frac{\dot{s}}{k} = (1-\alpha)\ln\alpha^{-1} + \frac{1}{3(1+\alpha)}[\Delta H - W_m + \alpha^2 D(q\|l)], \quad (14)$$

where the first term $(1-\alpha)\ln\alpha^{-1}$ is due to friction between the demon and the tape.

We now analyze the efficiency of the demon. First we consider the case where $q > 1/2$ and $l > q$, so that the demon operates as a machine. If we maximize the power $\dot{w}_m$ with respect to q, i.e. the load f, for fixed l and α we find that the maximal power $\dot{w}_m^*$ is reached at q^* given by the solution of the transcendental equation

$$q^* + (1-q^*)q^*\ln(q^*/(1-q^*)) = l, \quad (15)$$

independent of α. The efficiency at maximum power is defined as

$$\eta^*(\alpha, l) \equiv \dot{w}_m^*/(\dot{s}^* + \dot{w}_m^*), \quad (16)$$

where $\dot{s}^*$ is the entropy production at q^*. This efficiency depends on α and goes to zero for $\alpha \to 0$ because the dissipation term $(1-\alpha)\ln\alpha^{-1}$ dominates the entropy production. For $\alpha = 1$, the dissipation due to friction is zero, and the dissipation due to the refeeding of the tape becomes more relevant. More precisely, there is an optimal $\alpha^*(l)$, where the efficiency $\eta_{\max}^*(l) \equiv \eta^*(\alpha^*(l), l)$ is maximal. In Fig. 3, we plot $\alpha^*(l)$, which is always close to 1 and becomes closer to 1 as l tends to $1/2$. The efficiency $\eta_{\max}^*(l)$ increases as we approach $l = 1/2$, where it becomes $1/2$.

If $q > 1/2$ and $l < q$ the demon erases information with a rate $-\dot{h} = -k\Delta H/(3+\alpha)$. If we maximize $-\dot{h}$ with respect to l for fixed q and α, the maximum occurs for $l^\dagger = 1/2$. The efficiency at maximum "erasure rate" is then

$$\eta^\dagger(\alpha, q) \equiv -\dot{h}^\dagger/(\dot{s}^\dagger - \dot{h}^\dagger), \quad (17)$$

where $\dot{s}^\dagger$ is the entropy production for $l^\dagger = 1/2$. Again, as shown in Fig. 3, the optimal $\alpha^\dagger(q)$ is close to 1 and approaches 1 for $q \to 1/2$. The efficiency $\eta_{\max}^\dagger(q)$ at $\alpha^\dagger(q)$ also increases as q approaches $1/2$, where it becomes $1/3$.

In the final part, we introduce the linear response theory for this demon, for which we go back to the chemical model. In linear response, the stationary fluxes take the form $J_m = \sum_{n=1}^{6} L_{mn}\mathcal{F}_n$, where the affinities $\mathcal{F}$ are given in table 1 and the Onsager coefficients L_{mn} are defined as

$$L_{mn} \equiv \left.\frac{\partial J_m}{\partial \mathcal{F}_n}\right|_{\mathcal{F}=0}. \quad (18)$$

Calculating this coefficients we obtain for the matrix L

$$\frac{k}{72}\begin{pmatrix} 72 & 12 & 12 & -12 & -12 & 0 \\ 12 & 5 & -1 & -2 & -2 & 3 \\ 12 & -1 & 5 & -2 & -2 & -3 \\ -12 & -2 & -2 & 5 & -1 & 3 \\ -12 & -2 & -2 & -1 & 5 & -3 \\ 0 & 3 & -3 & 3 & -3 & 6 \end{pmatrix}. \quad (19)$$
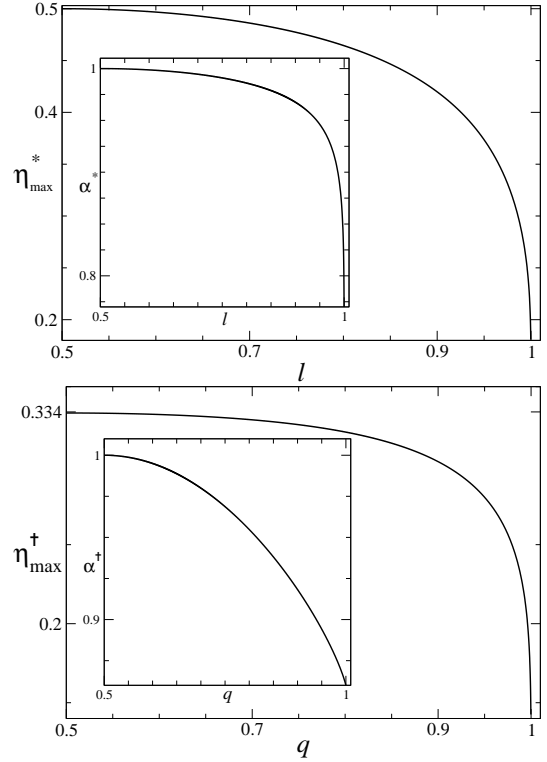

Fig. 3: Upper panel: the efficiency $\eta_{\max}^*$ as a function of l, when the model operates as a machine. Lower panel: the efficiency $\eta_{\max}^\dagger$ as a function of q, when the model is an eraser. In both cases the optimal α is shown in the inset.

Note that this matrix is symmetric, in agreement with the Onsager reciprocity relations $L_{mn} = L_{nm}$. It is obtained by considering all four chemical potential as independent variables. Using the restriction on the chemical potentials, within linear response we have $\mathcal{F}_2 = -\mathcal{F}_3$ and $\mathcal{F}_4 = -\mathcal{F}_5$. In addition, with the self-consistency condition (11) we obtain $\mathcal{F}_6 = -2\mathcal{F}_4$. Thus, there are only three independent affinities, which we define as $\tilde{\mathcal{F}}_1 \equiv \mathcal{F}_1$, $\tilde{\mathcal{F}}_2 \equiv \mathcal{F}_2$ and $\tilde{\mathcal{F}}_3 \equiv \mathcal{F}_6$. Moreover, the associated currents are $\tilde{J}_1 = J_1$, $\tilde{J}_2 = J_2 - J_3$, and $\tilde{J}_3 = J_6$.

The linear response entropy production is given by $\dot{s} = \sum_{m=1}^{6}\sum_{n=1}^{6} L_{mn}\mathcal{F}_m\mathcal{F}_n$. Using the three independent affinities it can be rewritten as $\dot{s} = \sum_{m=1}^{3}\sum_{n=1}^{3}\tilde{L}_{mn}\tilde{\mathcal{F}}_m\tilde{\mathcal{F}}_n$, where the the matrix $\tilde{L}_{mn} \equiv \left.\frac{\partial \tilde{J}_m}{\partial \tilde{\mathcal{F}}_n}\right|_{\tilde{\mathcal{F}}=0}$ is given by

$$\frac{k}{72}\begin{pmatrix} 72 & 0 & 0 \\ 0 & 12 & 6 \\ 0 & 6 & 3 \end{pmatrix}. \quad (20)$$

Note that in order to calculate this matrix the relations between the chemical potentials and the self-consistency condition have to be implemented before taking the derivatives. The explicit form of the entropy production is then

$$\dot{s}/k = (\ln\alpha^{-1})^2 + (2/3)(l-1/2)[(l-1/2) - (q-1/2)]$$
$$- (2/3)(q-1/2)[(l-1/2) - (q-1/2)]. \quad (21)$$

The second term in the right hand side is given by $\tilde{\mathcal{F}}_2\tilde{J}_2$, and is related to the dissipation term $\frac{k}{3(1+\alpha)}[\Delta H + \alpha^2 D(q||l)]$ in (14), whereas the third term is $\tilde{\mathcal{F}}_3\tilde{J}_3$, which corresponds to the mechanical work.

Using these results we can prove bounds on the efficiency at maximum power in the linear response regime. For $q > 1/2$ and $l > q$, where the demon operates as a machine, the power for fixed l and α is maximal at $q^* = (l + 1/2)/2$. Calculating the power and the entropy production at q^*, we obtain with the definition (16) the following efficiency at maximum power

$$\eta^*(\alpha,l) = \frac{1}{2}\left(\frac{1}{3[\ln(\alpha^{-1})/(l-1/2)]^2+1}\right) \leq \frac{1}{2}. \quad (22)$$

When the demon operates as an eraser, noting that within linear response $\Delta H = 2[(l - 1/2)^2 - (q - 1/2)^2]$, we obtain that the erasure rate for fixed q and α is maximal at $l^\dagger = 1/2$. Calculating the linear response theory entropy production at $l^\dagger$ we obtain

$$\eta^\dagger(\alpha,q) = \frac{1}{3}\left(\frac{1}{[\ln(\alpha^{-1})/(q-1/2)]^2+1}\right) \leq \frac{1}{3}. \quad (23)$$

Hence, when the demon operates as a machine the efficiency at maximum power is bounded by the well known Curzon-Ahlborn result $1/2$, valid for strongly coupled machines [25, 28]. However, when the demon operates as an eraser the efficiency at maximum erasure rate is bounded by $1/3$.

In conclusion, our stochastic fully time-continuous version of a Maxwell's demon contains thermal and mechanical equilibrium as a possible state. It thus allows to develop a systematic linear response theory and to identify Onsager coefficients. As for ordinary machines, we find that efficiency at maximum power is bounded by $1/2$. Operating as an eraser, however, the efficiency at maximum erasure rate is bounded by $1/3$. These linear response results can be extended to the non-linear regime where we recover in the irreversible limit the inequality between mechanical power and information flow found in [14]. The derivation of our results was facilitated by exploiting an analogy between this demon and an enzyme transforming and transporting molecules between two compartments. Exploring whether and how these results can contribute to a unified theory of autonomous machines that includes the aspect of information processing at finite temperature explicitly remains a challenging perspective for future work.

$$* * *$$

REFERENCES

[1] LEFF H. S. and REX A. F., *Maxwell's Demon : Entropy, Classical and Quantum Information, Computing* (IOP, Bristol and Philadelphia) 2003.
[2] MARUYAMA K., NORI F. and VEDRAL V., *Rev. Mod. Phys.*, **81** (2009) 1.
[3] CAO F. J., DINIS L. and PARRONDO J. M. R., *Phys. Rev. Lett.*, **93** (2004) 040603.
[4] ANDRIEUX D. and GASPARD P., *Proc. Natl. Acad. Sci. U.S.A.*, **105** (2008) 9516.
[5] GRANGER L. and KANTZ H., *Phys. Rev. E*, **84** (2011) 061110.
[6] VAIKUNTANATHAN S. and JARZYNSKI C., *J. Chem. Phys.*, **134** (2011) 054107.
[7] AVERIN D. V., MÖTTÖNEN M. and PEKOLA J. P., *Phys. Rev. B*, **84** (2011) 245448.
[8] HOROWITZ J. M. and PARRONDO J. M. R., *EPL*, **95** (2011) 10005.
[9] HOROWITZ J. M. and PARRONDO J. M. R., *New J. Phys.*, **13** (2011) 123019.
[10] ABREU D. and SEIFERT U., *EPL*, **94** (2011) 10001.
[11] BAUER M., ABREU D. and SEIFERT U., *J. Phys. A Math. Theor.*, **45** (2012) 162001.
[12] ESPOSITO M. and SCHALLER G., *EPL*, **99** (2012) 30003.
[13] KISH L. B. and GRANQVIST C. G., *EPL*, **98** (2012) 68001.
[14] MANDAL D. and JARZYNSKI C., *Proc. Natl. Acad. Sci. U.S.A.*, **109** (2012) 11641.
[15] TOUCHETTE H. and LLOYD S., *Phys. Rev. Lett.*, **84** (2000) 1156.
[16] TOUCHETTE H. and LLOYD S., *Physica A*, **331** (2004) 140.
[17] CAO F. J. and FEITO M., *Phys. Rev. E*, **79** (2009) 041118.
[18] SAGAWA T. and UEDA M., *Phys. Rev. Lett.*, **104** (2010) 090602.
[19] HOROWITZ J. M. and VAIKUNTANATHAN S., *Phys. Rev. E*, **82** (2010) 061120.
[20] ABREU D. and SEIFERT U., *Phys. Rev. Lett.*, **108** (2012) 030601.
[21] SAGAWA T. and UEDA M., *Phys. Rev. E*, **85** (2012) 021104.
[22] SAGAWA T. and UEDA M., *Phys. Rev. Lett.*, **109** (2012) 180602.
[23] TOYABE S., SAGAWA T., UEDA M., MUNEYUKI E. and SANO M., *Nature Phys.*, **6** (2010) 988.
[24] BÉRUT A., ARAKELYAN A., PETROSYAN A., CILIBERTO S., DILLENSCHNEIDER R. and LUTZ E., *Nature*, **483** (2012) 187.
[25] SEIFERT U., *Rep. Prog. Phys.*, **75** (2012) 126001.
[26] HOROWITZ J., SAGAWA T. and PARRONDO J. M. R., *arXiv:1210.6448*, (2012) .
[27] COVER T. M. and THOMAS J. A., *Elements of information theory* (Wiley, Hoboken, NJ, and Canada) 2006.
[28] ESPOSITO M., LINDENBERG K. and VAN DEN BROECK C., *Phys. Rev. Lett.*, **102** (2009) 130602.